Review: "A multi-instrument fuzzy logic boundary-layer top detection algorithm"

The authors' reply clarifies some of my comments. However, there still some confusing statements/definitions/discussions for me as listed below.

Comments:

1. While I concur that leveraging the synergy of multi-remote sensing boundary layer profiling measurements could enhance PBLH estimations by providing more constraining information, I find the authors' claim that 'the proposed method enables accurate boundary-layer height estimation both during daytime and nocturnal conditions' unconvincing based on the comparisons in Figures 8-11. The PBLHs estimated by remote-sensing (Fuzzy-logic) are consistently lower than those derived from radiosondes for PBLHs > 500 m or during afternoon convective boundary layer conditions (Figure 10). Why or what causes the underestimations from Fussy-logic PBLHs?
2. Based on Figure 2, it appears that the second-generation step yields even lower PBLHs than the first-generation step. This suggests that the second-generation step further underestimates PBLHs under conditions of a convective boundary layer. Is that correct?
3. The comparisons between Fuzzy-logic and radiosonde PBLHs do not seem as robust as those shown in previous studies that exclusively used Doppler lidar measurements, such as those in Tucker et al. (2009) and Krishnamurthy et al. (2021). Could you provide the correlation coefficients between Fuzzy-logic PBLHs and Radiosonde PBLHs as depicted in Figures 8, 9, and 11?

   References:

   Tucker, S. C., C. J. Senff, A. M. Weickmann, W. A. Brewer, R. M. Banta, S. P. Sandberg, D. C. Law, and R. M. Hardesty, 2009: Doppler Lidar Estimation of Mixing Height Using Turbulence, Shear, and Aerosol Profiles. *J. Atmos. Oceanic Technol.*, **26**, 673–688, https://doi.org/10.1175/2008JTECHA1157.1.

   Krishnamurthy, R., Newsom, R. K., Berg, L. K., Xiao, H., Ma, P.-L., and Turner, D. D.: On the estimation of boundary layer heights: a machine learning approach, Atmos. Meas. Tech., 14, 4403–4424, https://doi.org/10.5194/amt-14-4403-2021, 2021.
4. The authors have noted that in the first-generation step, the temperature inversion height was used as an input. Could you clarify how the temperature inversion height is derived? What degree of temperature difference is considered to constitute an inversion? The term 'temperature inversion height' is introduced in line 139 without an explanation of how it is obtained. In the response, the authors claim, *'We have never stated that inversion height is derived from the first-generation step*.' If so, could you specify when the 'inversion height' is derived?

5. Typically, a figure caption should provide a brief explanation of what the figure represents, such as the observations or variables depicted in Figure 2. This would provide readers with basic information, and I don't believe it would be as 'extensive' as the authors suggested in their response. They could relocate the detailed discussions or illustrations of the figure to the main body of the text.
6. Despite several rounds of discussion, the precise definitions of 'buoyancy-driven processes' and 'mechanically-driven processes' remain unclear.
7. The authors replied that '*if dz = the spacing between any two levels in a profile, and we know that dz varies throughout the profile, 300 is the average of all dz values in the profile. The minimum dz that occurs in the lowest levels of the atmosphere is on the order of 10s of meters. The maximum dz that occurs higher up in the atmosphere is on the order of 100s of meters.*' If a group of numbers have a minimum of 10 and maximum of 100, how can the average to be 300?
8. As for the comparison of high- vs. low-resolution sounding data, the authors state, '*Analyzing the methods separately highlights that using high-resolution sounding data tends to lead to lower median BL height, but the ranges and interquartile range values do not change much between coarse and high resolution datasets. There is no indication from this analysis that using high resolution data necessarily yields more accurate BL height values*'. It appears that the authors are suggesting that because 'the coarse and high resolution datasets have similar PBLH ranges and interquartile range values', therefore, '*There is no indication from this analysis that using high resolution data necessarily yields more accurate BL height values'.* I don't believe this is a solid conclusion. For instance, under an extreme condition where the coarse sounding data produce identical PBLHs, resulting in ranges and interquartile range values of zero, it certainly doesn't imply that coarse sounding data provides accurate PBLH estimations.